

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

Curriculum-Based Measurement

Buros-Nebraska Series on Measurement and  
Testing

---

1993

## 6. CBA: An Assessment of Its Current Status and Prognosis for Its Future

Mark R. Shinn  
*University of Oregon*

Roland H. Good III  
*University of Oregon*, [Roland\\_Good@ccmail.uoregon.edu](mailto:Roland_Good@ccmail.uoregon.edu)

Follow this and additional works at: <https://digitalcommons.unl.edu/buroscurriculum>



Part of the [Curriculum and Instruction Commons](#), and the [Educational Assessment, Evaluation, and Research Commons](#)

---

Shinn, Mark R. and Good, Roland H. III, "6. CBA: An Assessment of Its Current Status and Prognosis for Its Future" (1993). *Curriculum-Based Measurement*. 8.

<https://digitalcommons.unl.edu/buroscurriculum/8>

This Article is brought to you for free and open access by the Buros-Nebraska Series on Measurement and Testing at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Curriculum-Based Measurement by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

## CBA: An Assessment of Its Current Status and Prognosis for Its Future

Mark R. Shinn and Roland H. Good, III

*University of Oregon*

The very fact that curriculum-based assessment (CBA) forms the basis of a topic-driven conference at the center of American educational and psychological measurement (i.e., the Buros Institute) is testimony that the strategies are receiving a substantial amount of professional attention. Although debate continues regarding to whom and when the term curriculum-based assessment should be ascribed (Coulter, 1988), without question, its prominence has grown considerably in the last 10 years. Within the last 5 years, school psychology and special education have seen their flagship journals, *School Psychology Review* and *Exceptional Children*, devote special volumes to CBA. National organizations such as the National Association of School Psychologists (NASP) and the National Coalition of Advocates for Children (NCAS) have encouraged the use of CBA for decision making with handicapped students (NASP /

---

Authors' Notes. The development of this chapter was supported in part by Grant No. 8029D80051-90 from the US Department of Education, Special Education Programs, to provide leadership training in curriculum-based assessment. The views expressed within this chapter are not necessarily those of the USDOE.

NCAS, 1985). Interpretations of recent litigation also have been construed to suggest use of CBA strategies (Reschly, Kicklighter, & McKee, 1988a; Reschly, Kicklighter, & McKee, 1988b; Reschly, Kicklighter, & McKee, 1988c).

This chapter seeks to examine CBA's future as an assessment strategy from a perspective of school systems change (Sarason, 1982) and adoption of technological innovations (Rogers, 1983). To understand the school-change process, Hall and Hord (1984) maintain that change agents must consider the perspective of the *implementors* of the innovation. Using what they call a Concerns-Based Adoption Model (CBAM), Hall and Hord (1984) propose that implementors' concerns about change progress through a sequence of seven stages: (a) awareness, (b) informational, (c) personal, (d) management, (e) consequence, (f) collaboration, and (g) refocusing. An individual's concerns about innovation are not confined to any one stage, however. The seven concerns are divided into four general categories. Awareness is categorized as an *unrelated concern*, where the implementor generally is only somewhat cognizant of the innovation. Informational and personal concerns are *self concerns*, where the implementors' reactions are centered primarily on how the innovation affects them. Management is a *task concern*, where consideration is given to how best to use the innovation. Consequence, collaboration, and refocusing are *impact concerns*, where attention is shifted to the potential effects of the innovation on clients.

Each stage of Hall and Hord's CBAM model requires a different approach to influencing and facilitating the change process. At best, we believe the field of education, and more specifically special education and school psychology, is currently at the awareness and informational stages with respect to the implementation of CBA. Professionals are being exposed to CBA and are gathering information. We believe that an analysis of the future of CBA will require us to examine first the extent of professionals' knowledge regarding CBA. We will accomplish this task in two ways. First, we will identify briefly the major innovators in CBA and where their information is being disseminated. Second, we will analyze the major critiques of CBA (Lentz & Shapiro, 1986; Lombard, 1988a; Lombard, 1988b; Taylor, Willits, & Richards, 1988) under the premise that one gains an understanding of *what* is being communicated by *how accurately* it is described by others than the innovators themselves.

Before we can consider widespread adoption of CBA procedures, we must move beyond the informational stage of the CBAM model. To accomplish this movement, we need to analyze the information being communicated about CBA to ensure its accuracy. This chapter presents key discriminations that we believe implementors must make for

informational needs to be satisfied within the CBAM model. Additionally, given adoption of a scientist-practitioner model, we will identify the pieces of information and data that must be generated to validate empirically the various CBA strategies. This chapter therefore concludes with our analysis of future research needs.

**Table 1**

A sampling of articles on curriculum-based assessment published in refereed journals through 1989.

Journal	Authors
<i>American Educational Research Journal</i>	Fuchs, Deno, & Mirkin, 1984
<i>Diagnostique</i>	Fuchs, Deno, & Marston, 1983; Marston, Fuchs, & Deno, 1986
<i>Exceptional Children</i>	Blankenship, 1985; Deno, 1985; Fuchs, Fuchs, & Deno, 1985; Galagan, 1985; Gickling, & Thompson, 1985; Marston, & Magnusson, 1985; Rosenfield, & Robinson, 1985; Tucker, 1985
<i>Focus on Exceptional Children</i>	Deno, & Fuchs, 1987
<i>Journal of Behavioral Assessment</i>	Good & Shinn, in press; Mirkin, Deno, Tindal, & Kuehnle, 1982
<i>Journal of Educational Research</i>	Fuchs, Fuchs, & Tindal, 1986b; Tindal, et al., 1985
<i>Journal of Learning Disabilities</i>	Shinn, Ysseldyke, Deno, & Tindal, 1986
<i>The Journal of Special Education</i>	Fuchs, & Fuchs, 1986b; Marston, 1988
<i>Journal of Special Education Technology</i>	Fuchs, Deno, & Mirkin, 1983
<i>Learning Disability Quarterly</i>	Deno, Wesson, & King, 1984b; Shinn, Tindal, Spira, & Marston, 1987; Wesson, King, & Deno, 1984
<i>Professional School Psychology</i>	Shinn, Tindal, & Stein, 1988
<i>Reading Research Quarterly</i>	Fuchs, Fuchs, & Deno, 1982
<i>Remedial and Special Education</i>	Fuchs, 1986; Fuchs, & Fuchs, 1984; Shinn, & Marston, 1985; Tindal, Shinn, & Germann, 1987
<i>School Psychology Review</i>	Deno, 1986; Howell, 1986; Neisworth, & Bagnato, 1986; Shinn, 1986; Shinn, 1988; Shinn, Rosenfield, & Knutson, 1989
<i>TEACHING Exceptional Children</i>	Deno, Mirkin, & Wesson, 1984a; Wesson, 1987



## ANALYSIS OF THE PUBLISHED LITERATURE

### Analysis of the CBA Published Literature by Its Creators

As of January 1, 1990, over 100 articles, book chapters, or books have been published investigating or describing the use of CBA strategies (for a partial listing, contact the authors). The publication channels have included, but are not limited to, all the major special education journals and most school psychology journals. With the exception of a limited set of journals such as the *American Educational Research Journal* and the *Journal of Behavioral Assessment*, few articles about CBA have been published outside of these professional domains. A sampling of journals and prominent CBA authors is presented in Table 1.

Journal articles are supplemented by an increasing number of books, including ones by Hargis (1987); Idol, Nevin, and Paolucci-Whitcomb (1986); Howell (Howell & Kaplan, 1980; Howell & Morehead, 1987); Bagnato, Neisworth, & Munson (1989); Salvia and Hughes (1989); and Shinn (1989a), as well as training monographs/materials by Gickling and Havertape (1981) disseminated by the National Association of School Psychologists.

### Analysis of the CBA Published Critiques

The authors listed in Table 1 account for more than 95% of the research and scholarly articles written about CBA. An exhaustive review process failed to identify many articles written about CBA by persons other than these, although a number of resources (e.g., Will, 1986; 1989) mentioned CBA as a positive strategy. Among the eight articles that provided more than a cursory recommendation about the use of CBA, five were published in refereed journals (Reschly, 1988; Reschly et al., 1988a; Reschly et al., 1988b; Reschly et al., 1988c; Taylor, Willits, & Richards, 1988), one was a book chapter (Lentz, 1988), one was an article published in the newsletter of the National Association of School Psychologists (Lombard, 1988a) that was based on a paper presented at a state conference (Lombard, 1988b), and one was a letter to the editor in the NASP newsletter (Coates, 1989).

In an article describing the future of school psychology, Reschly (1988) proclaimed CBA as one of the most important new competencies required for school psychologists in alternative service delivery systems. He described CBA as educational assessment tools derived from a behavioral assessment paradigm where behavior is measured directly in the natural (i.e., classroom) environment. CBA was presented as a precise methodology for "measuring target behavior, monitoring

progress, and assessing outcomes" (p. 471). Further, Reschly suggested that CBA facilitates instruction on relevant skills. His description of CBA concluded with two caveats. First, professionals need specific training on CBA, as it is not a simple methodology. Second, to avoid misconceptions, it must be remembered that CBA is *not* (our emphasis) an intervention. Reschly, Kicklighter, and McKee (1988a; 1988b; 1988c) also commented favorably on CBA in a series of articles summarizing federal court cases on assessment and disproportionate placements in special education. In reviewing the rulings from the Marshall et al. vs. Georgia case (1984), they concluded that "the kind of assessment fostered by the Marshall Court is what has been called curriculum-based assessment. . . . CBA and other direct measures of functioning are preferable because the (assessment) results are related to interventions beneficial to the individual" (p. 20).

A more extensive critique of CBA was provided by Taylor, Willits, and Richards (1988) in an article published in *Diagnostique*. In describing CBA, Taylor et al. proposed that it was not really a new concept, and in fact, simply "formalized a long standing practice" (p. 15). CBA was essentially criterion-referenced testing (CRT) where curricular objectives were operationalized into tests and cutting scores were used to determine mastery. Many of Taylor et al.'s criticisms therefore centered on the weaknesses of CRTs. Foremost among the criticisms was that of the limited utility of CBA in assessment and decision-making practices. As stated by Taylor et al., "It is clear that CRTs alone are not sufficient to serve the many and diverse purposes of assessment. Consequently, it is doubtful that CBA will either" (p. 15). As a result of their purported limited utility, Taylor et al. recommended that CBA should be used only as a *supplemental* assessment strategy and should not supplant traditional assessment methods.

Taylor et al. went on to detail a number of other concerns about CBA. Among them, concern was expressed that the use of CBA for writing Individualized Education Plan (IEP) objectives would be a "loss of the individual" and that the content of the CBA test would dictate the content of instruction. Taylor et al. also noted concerns that the assessment procedures derived from a curriculum could not be valid if the curriculum was not valid. We assume that *valid* in the last use was used as a synonym for *effective*. Relatedly, concern was expressed that a curriculum (and thus, CBA) may not reflect the needs of special education students. Other criticisms centered on CBA's use of local norms and the technical adequacy (i.e., reliability, validity) of the measures themselves. With respect to the former, Taylor et al. argued

that the local norms developed for CBA would be difficult to interpret and would result in special education students' change of eligibility, depending on the school system in which they were enrolled. Taylor et al. (1988) concluded their critique of CBA with its positive use only in the following set of conditions:

1. If the curriculum on which the CBA is based is valid.
2. If the curriculum on which the CBA is based represents the needs of the special education student.
3. If the CBA instrument can be developed to yield reliable and valid results.
4. If limitations are acknowledged or additional research is conducted regarding the curricular areas for which CBA is appropriate.
5. If limitations are acknowledged regarding the use of CBA as a comprehensive assessment approach.
6. If careful attention is given to properly training users of CBA.

In his chapter on direct observation and measurement of academic behavior, Lentz (1988) describes CBA as employing direct measures of academic behavior that are essential to the resolution of academic problems in the classroom. CBA is seen as oriented to the determination of special education eligibility, setting individual educational plan (IEP) goals, and monitoring progress using procedures that were designed to offset the problems with "norm-based achievement tests" (p. 84). Tests are short-duration probes that assess the academic skills taught within the classroom using stimulus materials from the instructional curricula. In contrast to criterion-referenced tests, CBA procedures are used in a repeated fashion. While noting these strengths, Lentz provided a number of criticisms of CBA from a behavioral perspective. Among the criticisms was his contention that CBA research was conducted out of a nonbehavioral, psychometric approach where probes are high-inference measures about global constructs. Lentz also took issue with the use of CBA probes for problem identification/screening as a process that "does not fit a behavioral model very well" (p. 103). Finally, he criticized CBA for its lack of utility in specifying which treatments will work. As stated by Lentz (1988), "It seems clear that CBA probe data cannot be used unilaterally to predict success of interventions" (p. 106).

The most critical review of CBA was written by Lombard (1988a). In critiquing one type of CBA, curriculum-based measurement (CBM), he asserted that it had not lived up to its promise as a "new and improved paradigm to meet special education students needs" (p. 20).

Lombard's major criticisms fell into two major categories: (a) the components of what was measured and (b) the use of the measures for purposes in making special education eligibility decisions. His concerns about what comprised the CBM probes were similar to those cited by Taylor et al. (1988), including curriculum bias, speed effects, effects of students' attentional and psychomotor deficits on their scores, and what he referred to as the tests' limited behavior sampling. Lombard's concerns about CBA were directly counter to the Reschly et al. interpretation of the *Marshall* (1984) court case. Lombard expressed concern that CBA strategies were both discriminatory towards minorities and would redefine the special education population by placing low-achieving, not-truly-handicapped students in special education. Further, he stated that the use of CBA has allowed the general education system to "short-cut" the requirements of PL 94-142.

The final critique by non-CBA authors was that of Coates (1989). In his brief but succinct commentary, Coates praised curriculum-based assessment as an exciting new measurement technology. However, he also raised concerns about the apparent assumption of many CBA proponents that standardized norm-referenced tests have no usefulness beyond placement decisions and the notion that norm-referenced testing and CBA are antagonistic, as well as concerns about the validity of CBA reading measures.

### CBA Informational Needs for Educators

How does one reconcile the differences in interpretations and criticisms of CBA by authors such as Reschly, Taylor, Lentz, Lombard, and Coates? If Hall and Hord's concerns-based adoption model is employed, what current informational needs are suggested to allay personal concerns and facilitate implementation of this innovative technology? Based on our analysis and knowledge of the published CBA references and the criticisms of CBA, we see the need to engage in a series of *discriminations* within the existent literature, including distinguishing between (a) assessment terms, (b) assessment decisions, (c) different models of CBA, (d) assessment paradigms, and (e) CBA-based changes and the change process itself.

### Discriminating Between Assessment Terms

The easiest discrimination that can be made within the existent literature on CBA is to clarify the terms that are used to describe both CBA and other measurement tools. We have observed the terms *assessment*, *standardized*, *norm referenced*, *criterion referenced*, *informal*,

*formal*, and *published* to be bandied about almost casually, and often interchangeably. We propose that all authors increase the precision of the language used to describe various measurement terms. As two cases in point, consider the term *norm referenced* as used by Coates (1989) and Lentz (1988). Coates asserted that CBA is, in a sense, against "standardized norm-referenced" tests. Lentz described CBA as a system developed to overcome problems with "norm-based" achievement tests. In both cases, the authors are referring to commercially available, norm-referenced achievement tests. The key term is *commercially available*, not *standardized* or *norm referenced*. CBA can be standardized (i.e., administered and scored in a prescribed, replicable manner) and can be used in a norm-referenced manner where a specific student's score is compared to a normative sample (Shinn, 1989b). The use of terms *informal* and *formal*, with the former implying either nonstandardized and/or not commercially available and the latter implying standardized and/or commercially available and/or norm referenced, contribute little information and less ambiguous terms are available. We believe the salient features of academic assessment can be described using the following terms and definitions:

1. *Standardized*: A test that is administered and scored in a specified, replicable manner.
2. *Nonstandardized*: A procedure for collecting data that is idiosyncratic to the examiner, with results that may have little generality across individuals and time.
3. *Commercially available*: A test or procedure that is produced by a publisher.
4. *Norm referenced*: A test that has interpretive metric(s) derived from a comparison group.
5. *Criterion referenced*: A test that has items derived from an identified instructional domain, with interpretive metric(s) derived rationally (i.e., without sampling from a group of students).
6. *Individually referenced*: A test that has items derived from an identified, finite instructional domain, with interpretive metric(s) derived by comparing the student's score to his or her previous scores over time.

All tests are standardized. Single terms thus may be used hierarchically. For example, a published, norm-referenced test (Woodcock Reading Mastery Test) implies, by definition, standardization. These distinctions can eliminate many confusions engendered by authors.

### Discriminating Between Models

The articles by Reschly (1988) and Taylor et al. (1988) provide clear evidence of the need to clarify that CBA is not a unified set of procedures or strategies. There is no *one* model of CBA. Although generally quite accurate in his description of CBA, Reschly (1988) errs in stating categorically that CBA is behavioral assessment applied to academic problems and that CBA is not an intervention. The specific accuracy of his statements is dependent upon which model of CBA is considered. Models of CBA range from those placing great reliance on a behavioral assessment paradigm (Deno, Mirkin, & Shinn, 1979; Knutson & Shinn, 1990; Shinn, Goodwin, & Habedank, 1989) to those that are decidedly nonbehavioral (Gickling & Havertape, 1981). With respect to the contention that CBA is not an intervention, it is important to note that *all* assessment, including CBA, is to some degree an intervention; data are derived to improve the functioning of the individual assessed. The degree to which CBA is or is not an intervention parallels the continuum of whether the CBA model is behavioral. The model of CBA represented by Deno currently represents the end of the continuum where it is less of an intervention. Gickling's model, on the other hand, represents the other end of the continuum, as it is almost exclusively an intervention strategy.

Errors of discrimination between models are made also by Taylor et al. (1988). As presented earlier, these authors consider CBA to be essentially criterion-referenced testing (CRT) where a curricular objective is identified and a test and mastery score are constructed to correspond to the domain that the objective represents. CBA is treated as synonymous with CRTs and Taylor et al. view it as having the same strengths and weaknesses. However, it is apparent from an examination of the reference list for the Taylor article that the authors are referring to five *different* models of CBA. We have classified the types of CBA model and authors in Table 2. Only *two* of these models, the Blankenship CBA-CRT and the Bagnato, Neisworth, and Munson preschool CBA model, could be characterized as CRTs. The other models are not based on traditional conceptions or definitions of criterion-referenced testing. Although all derive their testing items from the curriculum, the accuracy-based model of CBA, Curriculum-Based Measurement, and CBA for instructional design do not create CRTs for each curricular objective, nor do they establish mastery criteria on a rational basis.

**Table 2**

Classification of the different CBA citations characterized as one CBA model in Taylor, Willits, and Richards (1988) into different models of CBA.

Author(s)	CBA Model
Blankenship, 1985	Criterion-referenced CBA
Bursick, & Lessen, 1987	CBA for Instructional Design
Deno, 1985; Fuchs, & Fuchs, 1986b; Lombard, 1988a; Lombard, 1988b; Marston, & Magnusson, 1985; Shinn, 1988; Wesson, King, & Deno, 1984	Curriculum-Based Measurement
Coulter, 1985; Rosenfield, & Rubinson, 1985	Accuracy-Based CBA
Neisworth, & Bagnato, 1986	CBA for preschool assessment

A growing number of professional resources are available that provide information for professionals to discriminate between the differing models of CBA (Marston, 1989; Shinn, Rosenfield, & Knutson, 1989; Tindal, this volume). It is beyond the scope of this chapter to detail sufficiently the important differences among CBA models. Suffice it to say that it is critical to discriminate among models. Failure to do so increases the likelihood of misunderstandings by practitioners. According to Hall and Hord (1984), lack of good information will impair resolution of the self-concerns in the systems-change process. It is important to note that discriminating among models does not imply incompatibility. Shinn, Rosenfield, and Knutson (1989) have argued that although the CBA models differ in some important ways, they have the potential to fit together to form a coherent problem-solving educational assessment system. Without discriminating between models, however, practitioners run the risk of overgeneralizing. In particular, they may misinterpret criticisms of *one* specific CBA model as pertaining to *all* CBA procedures. Technical adequacy (i.e., reliability, validity) is a case in point. Taylor et al. (1988) raised concerns about the technical adequacy of CBA. A novice in CBA may interpret Taylor's statement to be applicable to all models of CBA when one model, Curriculum-Based Measurement (CBM), has extensive documentation of its technical adequacy.



By combining injudiciously those features of the various CBA models that are genuine weaknesses, that are undeveloped (e.g., secondary applications of CBM), or that are beyond the intended focus of the model (e.g., school-age applications of Bagnato, Neisworth, and Munson's Preschool CBA), critics and practitioners can create the educational equivalent of an Edsel: a measurement and decision-making system that is indefensible. Alternatively, we believe firmly that selecting and combining specific strengths from across CBA models in practice can generate the educational equivalent of a Mercedes-Benz.

### Discriminating Between Assessment Decisions

In general, most assessment practices suffer from a lack of distinguishing what decision is to be made with the data. Although the use/overuse of published, norm-referenced tests (PNTs) is most frequently the target of criticism in this regard (Salvia & Ysseldyke, 1987), CBA also suffers for similar reasons (Shinn, Rosenfield, & Knutson, 1989). There appears to be a high likelihood of overstating the utility of the data derived from any test. As a result, we witness the continued practice of trying to plan instructional programs from PNTs, despite a lack of data to suggest that they can be used for such purposes (Deno, 1986). Similarly, we see some models of CBA being described as a "do-it-all" approach *without data* to do so. In order to select the most appropriate assessment procedure, one must first ask, "What decision am I being asked to make?" The demands placed on an assessment device vary with the educational decision being made.

Regardless of the strategies used to derive student data, we believe that assessment practices will be improved only when viewed within a decision-making context. Salvia and Ysseldyke (1987) have provided one decision-making model where data are collected to facilitate screening, eligibility determination, intervention planning, pupil progress, and program evaluation decisions. Their heuristic provides a mechanism by which assessors can select strategies for collecting data to make decisions. In recent years, we have adopted a decision-making paradigm that closely approximates that of Salvia and Ysseldyke. Within a *problem-solving paradigm*, educational decisions are classified as problem identification, problem certification, exploring alternative solutions, evaluating solutions, and problem solution. The first four of the decisions correspond roughly to those of Salvia and Ysseldyke. When the last decision, problem solution, is added, one has a framework for making decisions about individual students that is less student centered and more situation centered than the Salvia and Ysseldyke



paradigm (for a more detailed discussion, see Shinn, Nolet, & Knutson, 1990). Within a problem-solving model, a problem is defined as a difference between what is expected and what occurs. Each step of the problem-solving model specifies a measurement strategy (the data to be collected) and an evaluation strategy (the decision to be made). The measurement and evaluation activities, as well as specific data collection strategies within the problem-solving model, are summarized in Table 3.

**Table 3**

Summary of Problem-Solving Model Decisions, Measurement Activities, and Evaluation Activities.

Problem-Solving Decision	Measurement Activities	Evaluation Activities	Specific Tasks
Problem Identification	Record Differences Between Expectations and Student Performance	Does a Discrepancy Exist?	Peer-Referenced Assessment
Problem Certification	Describe Severity of Discrepancy and Available Resources in Environment That Many Reduce Discrepancy	Are Additional Services Beyond Those Currently Available in the Typical Environment Needed?	Survey-Level Assessment & Evaluation of General Education Modifications
Exploring Solutions	Estimate Expected Student Gains and Available Alternative Resources	Which Intervention Will Be Implemented? What Are The Intervention's Goals	Write Long-Term Goals, Design Intervention Plan
Evaluating Solutions	Monitor Program Intervention, Student Progress	Is Program Effective, Is Student Making Progress?	Collect Data, Compare Actual & Expected Performance
Problem Solution	Record Differences Between Expectations and Student Performance	Are Additional Resources Still Needed To Reduce Discrepancy	Repeat Peer-Referenced Assessment

Adapted from S. Deno (1989). Curriculum-Based Measurement and Special Education Services: A Fundamental and Direct Relationship. In M.R. Shinn, (Ed.) Curriculum-Based Measurement: Assessing Special Children, (pp. 1-17). New York, NY: Guilford Press.

Problem Identification and Certification place a high reliance on norm-referenced data to operationalize the severity of the discrepancy between what occurs and expectations. However, norm-based assessment strategies are less than useful for Exploring and Evaluating Solutions. Failure to discriminate between the decisions to be made and the data to be collected can result in inappropriate and ineffective assessment practices. Given the considerable differences that exist between CBA models with respect to their evidence for decision-making utility, failure to make these discriminations is likely to be common and problematic.

#### Discriminating Between Assessment Paradigms: Current and Problem-Solving Educational Assessment Practices

A key discrimination that must be made in this discussion is between CBA as an assessment technique (i.e., CBA as another "test") and the *paradigm* used to select and evaluate assessment techniques. The problem is not just that CBA techniques provide different data to answer the questions schools ask. Instead, we suggest that CBA *may* address different questions based on different underlying assumptions and values; in other words, a different paradigm. We add the caveat *may* in that, with the exception of CBM, the assumptions and values underlying most models of CBA have yet to be made explicit. The assumptions, philosophical underpinnings, and values specified overtly for CBM (e.g., Deno, 1985; 1986; 1989) clearly demonstrate fidelity to a different educational assessment paradigm, of which CBM is an important, but not the sole, component (Deno, 1989; Knutson & Shinn, in press). Our discussion of paradigm shift will focus, therefore, on the CBM model of CBA and the problem-solving paradigm.

We suggest that discussions of the value and future of CBA occur at two levels of discourse: paradigm and procedure. At the paradigm level are the values, assumptions, and regularities of current practice that generate the criteria by which we evaluate the adequacy of assessment techniques. At the procedure level is the evaluation of specific techniques or procedures with respect to established criteria. At the procedure level, we might ask, "How good is this assessment technique?" At the paradigm level, we might ask, "How will we know a good technique when we see one?" The paradigm/procedure distinction is crucial because decisions about quality are based on different types of information at each level. Technique questions are resolved empirically by comparing the extent to which alternative

procedures satisfy established assessment criteria (e.g., best reliability, strongest criterion-related validity). In contrast, paradigm conflicts are resolved on the basis of values and assumptions. What purpose should we be trying to accomplish with our assessments? *Why* do we want to accomplish this purpose? Data are involved only in more general terms, as broad strokes of the research brush regarding the empirical support for underlying assumptions.

The distinction between procedure and paradigm is important because educators are questioning both levels. With respect to the former, attention is focused on the technical adequacy of current CBA assessment techniques. With respect to the latter, professionals are struggling with the larger issue of what is the "best" or "right" way to make data-based decisions about students. We argue that the future of CBA is not dependent solely upon procedure but is entwined inextricably with resolving what is the best way to make assessment decisions. If CBM is used merely to accomplish the same goals and objectives as current techniques, based on the same underlying values and assumptions (i.e., as a *supplement* to current assessment techniques) with more content-valid devices, its future most likely will be short, and perhaps deservedly so. Practitioners already are experiencing difficulty keeping up with their caseloads and, most likely, additional time and assessment requirements will not be received with enthusiasm. Further, it is likely that assessment activities will continue to be used only for child-find, special-education-eligibility decisions and not to improve student outcomes.

Paradigm questions must be resolved before assessment procedures can be compared meaningfully. In order to evaluate the worth of an assessment technique, we must first determine the purposes we expect the procedure to accomplish and clarify the rationale for those purposes. Only when the goals and purposes of assessment are established can we compare how well alternative assessment procedures accomplish those goals. Comparing current and alternative paradigms requires clarification of the values, assumptions, purposes, and goals of assessment. Unfortunately, the current assessment paradigm is not well articulated, so discussions of paradigm shift are difficult.

To illustrate the implications of a paradigm shift, we have constructed our best understanding of the current assessment paradigm based on the *existing regularities* found in current practice. An examination of existing regularities is important from a systems-change perspective. Sarason (1982) asserts that for change in schools to take place, one must make two assumptions: (a) that the change is desirable according to some set of values and (b) that the intended outcomes are

clear. Sarason (1982) maintains that the implied outcomes of any change process are "changing the existing regularity, eliminating one or more of them, or producing new ones" (p. 96). A regularity is a programmatic or behavior occurrence that is supposed to have an intended outcome. It is often an unspoken, assumed belief that is not data based. One regularity cited by Sarason as an example is that generally children in this country go to school 5 days per week (Monday through Friday). Often, however, the intended outcome of the regularity itself (as in the previous example) may not be clear, and there frequently are no systems built into schools to ascertain the discrepancy between regularities and intended outcomes (Sarason, 1982).

Important existing regularities implicit in current assessment practices are compiled in Table 4. We do not assume this list of regularities to be exhaustive. These regularities impact both the information we attempt to obtain and the criteria by which we evaluate the quality of assessment techniques. Within the regularities are implied anticipated outcomes, social values, and methodological testing techniques.

**Table 4**

**Regularities Questioned by Implementation of Curriculum-Based Assessment as Embedded Within Problem Solving.**

1. Commercially available, norm-referenced tests are used mostly by psychologists in a diagnostic-prescriptive fashion to identify, in advance of treatment, the interventions that will be successful (Deno, 1986).
2. Commercially available, norm-referenced tests are used in a pre and post testing format, usually on a yearly basis by teachers to evaluate student progress and intervention effectiveness (Deno, 1986).
3. Group designs are used for making statements about the effects of individual student programs (Deno, 1986).
4. Instruction not individualized nor evaluated. An assumption is made that what works for one student works for all students (Deno, 1986).
5. Students only are examined intensively because they are the cause of academic problems (Alessi, 1989).
6. Handicapping conditions (e.g., learning disabilities) are identified by school psychologists' testing students using commercially available, norm-referenced tests (Heller, Holtzman, & Messick, 1982).
7. We don't evaluate alternative interventions (e.g., special education) systematically because we know they are effective and therefore do not need to be evaluated (Deno, 1986).

If existing regularities are to be changed, the outcomes, values, and assumptions must be examined explicitly to determine whether there is a defensible underlying paradigm and whether an alternative paradigm should be adopted. We have attempted to translate the existing regularities into 10 dimensions of assessment practices that embody a paradigm. These dimensions are presented in the first column of Table 5. In column 2, questions that allow one to determine the quality of the practice are provided for current assessment procedures. The evaluative questions in column 2 are drawn from classical test theory and standard instruction in tests and measurement. With regard to the *purpose* of assessment, for example, if the existing regularities are to group students by handicapping condition and to provide corresponding interventions (e.g., special education services) on the basis of published, norm-referenced tests, assessment techniques must discriminate among students reliably. Assessment techniques that generate spread or variability in individual performance consequently are judged more apropos than those that do not. The intended outcome presumably is to provide appropriate instruction and services to children grouped by their classification. That this is an assumption or belief and not a data-based outcome is evidenced by the pervasive difficulties documenting the efficacy of special education placement (Heller, Holtzman, & Messick, 1982), and the regularity that interventions are not evaluated systematically.

Earlier, we reported Sarason's (1982) contention that for school change to occur, it must be desirable based on some values. We believe that the professional values espoused by school psychology leaders (e.g., Bardon, 1988; Graden, Zins, & Curtis, 1988; Reschly, 1988), as well as the results of the most recent survey of NASP leaders and practitioners (Reschly, Genshaft, & Binder, 1987), suggest that change in the current assessment paradigm is desired. However, we also believe the outcomes of alternative assessment practices have not been examined with regard to the changes that would be required in existing regularities. Although widespread dissatisfaction has been expressed with the current assessment paradigm, there is as yet no consensus regarding the preferred alternative assessment paradigm.

The alternative assessment system we propose is problem-solving educational assessment. In this paradigm, the ecological educational assessment model described by Shapiro and Lentz (1985) and the behavioral assessment model described by Barlow, Hayes, and Nelson (1984) are integrated within the problem-solving sequence detailed by Deno (1989) presented earlier. The model also addresses advances in and extensions of classical test theory (e.g., Messick, 1989). Knutson

**Table 5**  
Different Questions Resulting from Paradigm Shift.

Dimension	Current Assessment Paradigm	Problem-Solving Paradigm
Purpose	Do assessment results spread out individuals facilitating classification/placement into groups?	Does assessment result in socially meaningful student outcomes for the individual?
Test Validity	Does the assessment device measure what it says it measures? Criterion-Related Validity: Does the test correlate with other tests purporting to measure the same thing? Construct Validity: Does the test display a stable factor structure?	Are the inferences and actions based on test scores adequate and appropriate (Messick, 1989)? Treatment Validity: Do decisions regarding target behaviors and treatments based on knowledge obtained from the assessment procedure result in better student outcomes than decisions based on alternative procedures (Hayes, Nelson, & Jarrett, 1986)?
Unit of Analysis	Groups: Probabilistic statements about individuals: Do students with similar assessment results most likely display similar characteristics?	Individuals: Does assessment show that this treatment is working for this student?
Time Line	Summative: Does the assessment indicate whether or not the intervention did work?	Formative: Does the assessment indicate whether or not the intervention is working?
Level of Inference	Does the assessment provide an indirect measure of an unobservable construct?	Does the assessment directly measure important target behaviors or skills?
Locus of the Problem	Does the assessment identify relevant student characteristics that contribute to problem etiology?	Does assessment identify relevant curriculum, instruction, and contextual factors contribute to problem solution?
Focus	Problem Certification: Does assessment accurately identify problems?	Problem Solution: Does the assessment accurately identify solutions?
Test Reliability	Are test scores stable over time? Are scores based on different behavior samples, obtained in different contexts/ settings consistent?	What factors account for the variability in student performance?
Context	Does the assessment provide a comparison with students receiving a nationally representative range of curriculum and instruction?	Does the assessment provide a comparison with students receiving comparable curriculum and instruction?
Dimension of Dependent Variable	Does the assessment provide information regarding the level of pupil performance?	Does the assessment provide information regarding the level of pupil performance and the slope of pupil progress?

and Shinn (in press) provide details as to how the problem-solving educational assessment paradigm is operationalized. The evaluative questions within a problem-solving model by dimension are presented in column 3 of Table 5.

We believe that for a paradigm shift to occur, we must contrast current and alternative assessment practices by their evaluative criteria within each dimension. The juxtaposition of assessment questions in Table 5 illustrates the fundamental and far-reaching differences in assessment resulting from a paradigm shift. To illustrate in more detail some of the fundamental differences between paradigms, we will contrast the use of intelligence tests in decision making with instructional problems and CBM within a problem-solving model. In current practice, intelligence tests are used frequently to assist in decision making about academic problems. A major purpose purportedly is to provide a prediction of future learning. Educators might want to evaluate, for example, "a student's ability to benefit from instruction." If inadequate learning or academic progress is predicted to occur as a result of the student's ability to benefit from the types of instruction available within general education settings (e.g., the student obtains an IQ below 70), the student customarily is identified as handicapped and special education services are recommended. With the instruction available in special education settings (i.e., individualized educational programs, modifications in the curriculum and instruction), the student is anticipated to make better academic progress.

Within the problem-solving paradigm outlined in Table 5, practice would differ substantially. A problem would be defined as a discrepancy between observed and expected behavior (Deno, 1989). Assessment would examine the student's academic progress in curricular material over time. If the level of student skills or the rate of student progress was not adequate, alternative interventions would be implemented and evaluated systematically. Interventions would include modifications of instruction, curriculum, and context variables not necessarily requiring special education services. Interventions resulting in improved academic progress would be maintained and modified. Perhaps more importantly, interventions that were ineffective for the individual student would be changed. From this perspective, the assessment of intellectual functioning does not contribute to educational decision making.

Using the assessment of intellectual functioning as an exemplar, the effects of a shift in paradigms are examined with respect to the dimensions of the dependent measure, the level of inference, the unit of analysis, and the context of assessment.

## DIMENSION OF THE DEPENDENT VARIABLE

A fundamental difference between assessment paradigms regards the dimension of the dependent variable. The current assessment paradigm features a one-dimensional view, stressing a static measure of the *level* of pupil skills only. The problem-solving paradigm includes a second dimension of performance—time—stressing a dynamic examination both of the level of pupil performance and the *slope* of pupil progress.

Considerable confusion exists in the professional literature between the assessment of slope and level. The level of pupil performance refers to the amount or extent of skills displayed by the student at one point in time. An estimate of level is obtained from one assessment. The slope of pupil progress refers to the rate at which the student is acquiring skills over time. Obtaining an estimate of slope requires repeated assessments of skill level over time and a procedure for summarizing the rate of change (Good & Shinn, 1990; Shinn, Good, & Stein, 1989). From a mathematical perspective, slope refers to the unit change in a dependent variable (Y) associated with a unit change in an independent variable (X):

$$\text{Slope} = \frac{Y_2 - Y_1}{X_2 - X_1} \quad (1)$$

Because intelligence tests typically are given in one sitting at one point in time, IQ tests are, by definition, measures of the level of pupil performance only. On this day, Billy obtained an IQ score of 85 on the WISC-R. This outcome means that on this day, on these tasks, and under these conditions, Billy displayed skills at a level of proficiency one standard deviation below the mean. In contrast, a problem-solving paradigm would stress the assessment of skills over time. Using CBM, for example, a student's skills would be assessed on a frequent, repeated basis, with the results plotted on a two-dimensional graph (time by level of skill). The slope of pupil progress then would be used to evaluate the efficacy of interventions and the need for alternative, potentially more intrusive, interventions.



## LEVEL OF INFERENCE

A second fundamental difference between assessment paradigms regards the level of inference entailed in decisions about individual students. In general, when compared to the problem-solving paradigm, the current assessment paradigm countenances a much higher level of inference as decisions are based on less observable constructs and less direct data, and entail more assumptions that are more difficult to substantiate or are less tenable (Kratochwill & Shapiro, 1988). As discussed previously, intelligence tests are measures of students' level of performance. However, they typically are used to make high inference statements about the future slope of pupil progress. When educators use an IQ test to determine a "student's ability to benefit from instruction," for example, they are making an inference about the slope of pupil progress. Substantial benefit corresponds to a steep slope; little benefit corresponds to a shallow slope. Indeed, many researchers define intelligence (i.e., ability or aptitude) in terms of slope. Carroll (1989), for example, notes that "aptitude is the name given to the variable or variables that determine the amount of time a student needs to learn a given task, unit of instruction or curriculum to an acceptable criterion of mastery under optimal conditions of instruction and student motivation" (p. 26). Thus, under fixed conditions of instruction, the student with higher ability would display the steeper slope of pupil progress (i.e., acquire skills in a shorter length of time). The correspondence of IQ to slope of pupil progress also is evident in the familiar formulation of the ratio IQ, the initial metric of intelligence tests. The ratio IQ is defined as:

$$\frac{\text{Ratio IQ}}{100} = \frac{\text{MA}}{\text{CA}} = \frac{\text{MA} - 0}{\text{CA} - 0} \quad (2)$$

Or, alternatively, as:

$$\frac{\text{Ratio IQ}}{100} = \frac{Y_2 - Y_1}{X_2 - X_1} \quad (3)$$

Thus, the ratio IQ represents the amount of change in intellectual skills associated with a unit change in time over the individual's entire life

span, or the slope of pupil progress on intellectual, problem-solving skills. Clearly, then, statements about the slope of pupil progress are one intended purpose of intellectual assessment.

The use of intelligence tests to make inferences about future learning is not altogether unreasonable. However, meaningful conclusions about the slope of pupil progress may be drawn from measures of the level of pupil performance (e.g., an intelligence test) only when appropriate assumptions are met. As illustrated in Figure 1, inferences about slope based on comparisons of level require four assumptions. First, students must be at the same level at the beginning of the relevant time period ( $\text{Time}_1$ ). For the ratio IQ, the implied time period begins at birth ( $\text{CA} = 0$ ) where, indeed, intellectual skills conceptually are identically 0. When shorter time periods are considered, as in the student's educational career or the current academic year, the assumption of equal entry levels is more difficult to support. If students display different entry-level skills, different final-level skills would *not* be indicative of differences in slope.

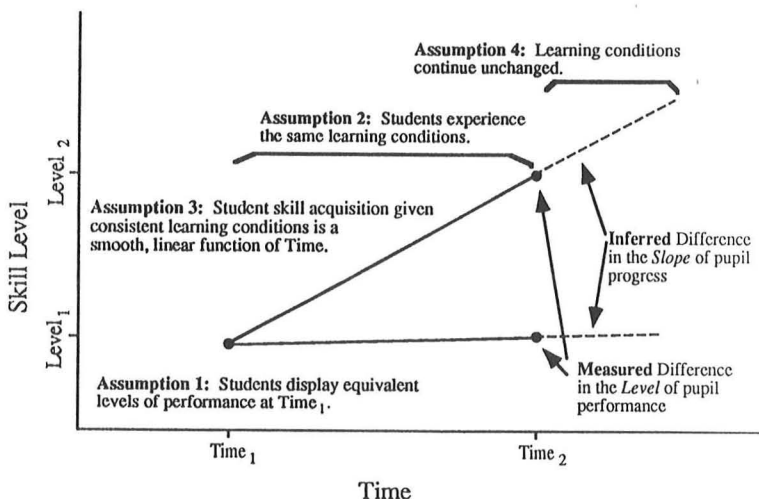


Figure 1. A graphic representation of the assumptions required to use a measure of level to infer slope.

The second assumption is that the students experienced identical learning conditions. To the extent that instructional conditions impact the slope of pupil progress (i.e., learning), different conditions would be confounded with differences in slope. Under disparate learning conditions, differences in the level of pupil skills could represent differences in the quality of instruction rather than a child characteristic.

The third assumption is that the acquisition of skills is a smooth, linear function of time, given consistent instruction. To the extent that the slope of pupil progress is sporadic or nonlinear, previous slope, especially over long time periods, would be less related to current or future slope. The fourth assumption is that learning conditions continue unchanged. A change in learning conditions would be expected to impact the slope of pupil progress, rendering inferences about current and future slope invalid.

Only when all four assumptions are tenable can inferences about the slope of pupil progress be made from differences in the level of pupil performance. When inferences about the slope of pupil progress in an academic content area are based on differences in the level of intelligence test performance, an additional, fifth assumption is necessary. This additional assumption is that the slope of pupil progress is consistent across skill areas. In particular, the slope of pupil progress on the tasks sampled by the intelligence test is assumed to be the same as the slope of pupil progress on academic skill measures, like oral reading fluency.

Clearly, making decisions about the slope of pupil progress based on intelligence test performance is a high-inference activity, requiring multiple assumptions that are difficult to assess and that vary in plausibility. It is no surprise that the few studies examining empirically the relationship between the slope of pupil progress and level of intellectual functioning have found little or no relationship (Bailey, 1981).

In contrast, a problem-solving educational assessment paradigm emphasizes a substantially lower level of inference. By assessing pupil progress *directly* in the skill area of interest, it is not necessary to assume that the slope of pupil progress is consistent across skill areas. By basing educational decisions on repeated measurements of academic skills over time, slope can be observed instead of inferred. It is not necessary to make extensive assumptions about instructional conditions and beginning skill levels. In addition, the conclusions drawn are at a much lower level of inference: At this time, under these instructional conditions, the slope of pupil progress was not adequate. Slope of pupil progress is not considered a student characteristic only, but is instead a combination of the student and the conditions of instruction. This approach requires a low-level assumption that the slope of pupil progress will continue unchanged in the absence of a change in instruction, curriculum, or conditions. However, a change in instructional conditions is not assumed to increase the slope of pupil progress. Instead, the slope of pupil progress following an intervention again is assessed.

## UNIT OF ANALYSIS

A third, fundamental difference between assessment paradigms regards the unit of analysis and interpretation. The assumptions required to make inferences about the slope of pupil progress based on measures of the level of intellectual functioning may be reasonable—for groups of students. In general, students are exposed to reasonably stable, homogeneous learning conditions (i.e., school) and enter school with roughly equivalent skills. Similarly, criterion-related validity studies repeatedly have demonstrated the relationship between intelligence test performance and academic achievement, again for groups of students. As a result, one can be completely confident that a *group* of students with low intelligence test scores will experience more difficulty in school than a group of students with high scores. Individual students with low scores, however, may *or may not* experience academic difficulty. Statements about individuals based on intelligence test scores are possible on a probabilistic basis only. With the relationship between academic achievement and intellectual functioning ranging between .60 and .80, students with low intelligence test scores will display substantial variability in academic performance. Some individuals will display quite high academic skills. Macmann, Barnett, Lombard, Belton-Kocher, and Sharpe (1989) provide an excellent illustration of this problem. They show that when two measures are correlated .80, and individuals are selected on the basis of extreme scores on one measure (i.e., 1.96 standard deviations below the mean), many cases will fall at or near the mean of the second measure.

From the perspective of the problem-solving paradigm, the question is *not* whether this individual student is a member of a group that, as a group, experiences academic difficulty. Instead, the question is whether *this individual student* is experiencing academic difficulty; the unit of analysis and interpretation is the individual.

## CONTEXT

The problem-solving paradigm differs substantially from the current assessment paradigm with respect to the role of context in the interpretation of assessment results. The context differences are epitomized by Taylor et al.'s (1988) arguments about local and national norms and the quality of the curriculum. These authors questioned, "How might CBA affect students performing at a satisfactory level within a school where the average student performance was considerably below average compared to other norms (national, state, or even

district)? The chances are that those students would not be identified for services even though they might need help" (p. 16). They also expressed concern that the school may not be using a "valid curriculum" (presumably one that is effective), and therefore that "CBA can be no better than the curriculum selected for instruction" (p. 17). We believe this point of view exemplifies most current assessment practices with respect to context, that a problem should reside solely within the student independent of context. Environmental expectations and characteristics, in terms of how other students perform or whether the curriculum is effective or ineffective, are not relevant to the identification of the problem. This position implies two potential outcomes: (a) that a student performing at a satisfactory level within a school where the average student performance is considerably below average compared to other (e.g., national) norms should be eligible for special education services, and (b) that a student performing considerably below expectations in his or her school but above other (national) norms should not be eligible for special education services. However, a focus on within-student pathology independent of context may be inconsistent both with best practices and with current practice.

It is crucial to examine more closely the implications of emphasizing within-child pathology independent of the context of the problem. Failure to consider context may result in untenable conclusions. In the first case, are we saying that identifying within-child pathology (e.g., learning disability or mental retardation) provides an acceptable amelioration for a dysfunctional system (e.g., ineffective curriculum)? Does this mean that the system can say five "Hail Marys," 10 "Our Fathers," place 15 children in special education, and receive absolution from the sins of its curriculum? In the second case, are we saying that we should do nothing because there is no "problem"?

In current practice, context effects on decision making regarding who receives special education services have been demonstrated empirically and repeatedly. For example, Singer, Palfrey, Butler, and Walker (1989) found in a recent study of five large school districts that districts "differed in the percentage of students they identified as handicapped, the frequency with which they used various labels, the criteria used to define groups, and the functional levels of students given the labels. Consistency was greatest for those labeled hearing impaired and, to a lesser extent, physically/multiply handicapped and weakest for those labeled mentally retarded and emotionally disturbed; results for those labeled speech impaired and learning disabled fell between these two extremes" (p. 278).

We agree that a problem exists when student performance is in the average range in the context of a school system that is substantially below average compared to national norms. We disagree that the problem is within the child or that placement in special education is the solution. Placing large numbers of students in special education will not change the fact that the school is severely below average compared to national norms and may not be providing an effective curriculum. Clearly, if the school or district is severely discrepant from national norms, the system has a schooling problem.

We also disagree that when student performance is below what is typical in a system that is above average compared to national norms, a significant problem does not exist. If the child is severely discrepant from expectations within the local context, the child may have a learning problem. For example, the child may exhibit low motivation, have poor attendance, display language difficulties, be receiving inappropriate or insufficient instruction, or be inappropriately placed in the curriculum. Individual interventions possibly necessitating special education services may be indicated.

Perhaps the future will hold a divided special education funding stream. One stream would fund services for individual students based on skills discrepant from local norms or expectations. A second stream would fund services for school systems or districts. A school district might be identified as *severely teaching disabled* (STD) based on performance discrepant from national norms or expectations. Special education services might include in-service training for teachers, improved curriculum materials, hiring incentives to attract and keep quality educators, and nutritional or early intervention programs for the community, among other possibilities.

### Distinguishing Between Changes in Practice as a Result of CBA and the Change Process Itself

One of our colleagues has self-titled a law about the change process (Stoner, personal communication, 1988). Stoner's Law goes something like this: When you ask someone to change, you are asking them to do more work. Asking people to do work often makes people angry. Therefore, when you ask people to change, you will make them angry. Under the best of circumstances, change will make only half the people involved angry; under the worst of circumstances, assume that change will make 95% of those involved angry. Introduction of CBA strategies in the schools is asking people to change. Whether CBA is an improvement to existing practices may be irrelevant when viewed in

the context of Stoner's Law. Attributes aside, we argue that we will need to discriminate implementation of CBA from the reactions to any change process. We can recall one particular circumstance where a school district was engaging in a general review of assessment and decision-making practices simultaneously with introduction of CBA. It was discovered by district personnel that no observations were being conducted prior to placement of students in programs for learning disabled students as required by state law. Resolution of the situation was interpreted (by teachers who had to conduct the observations) as being caused by CBA. In another district, we observed a school psychologist who was resistant to CBA centering his opposition on non-categorical placement, a school district practice that again was outside the direct effects of the implementation of CBA. Too often, changes in roles and responsibilities in general are often attributed to the innovation itself. Implementors should expect resistance to implementation and should work carefully to separate out the larger issues from those of implementing CBA.

## FUTURE KNOWLEDGE AND INFORMATION

We have taken the position that the evaluation of CBA should be based on an analysis of empirical outcomes, that useful assessment strategies should be documented to "work" in some way. An extensive body of research has been accumulated on CBA strategies in approximately 10 years. However, we are concerned that most of the empirical work has centered on CBM. Other CBA models have undergone little systematic inquiry. Many additional questions exist within CBM as well. We propose that the future information needs for successful implementation be examined in three separate areas: (a) establishment of technically adequate CBA measures, (b) use of the measures in decision making with students, and (c) research on implementation.

### Establishing Technically Adequate CBA Measures

Research on CBA measures must proceed in two interrelated areas. First, the pool of available measures with demonstrated technical adequacy must be increased. Second, CBA procedures must be identified for use with specific ranges of student populations (e.g., preschool, elementary, secondary).

*Technical adequacy.* We believe that CBA measures must meet professional standards for quality assessment devices if they are to be used for making important decisions with children. The major strategies



by which tests' quality is determined, a nomothetic, psychometric approach, or an idiographic, behavioral assessment approach, are merging in practice so that elements of both often are offered as evidence without contradiction (Barrios, 1988).

To date, only CBM researchers have undertaken extensive empirical studies of the technical aspects of their proposed instruments. CBM measures are constrained currently to the basic skills areas of reading, spelling, math, and written expression, with decreasing knowledge of technical properties in the respective order presented here. Although robust in their use with elementary-level and middle-school-level students with basic skill problems, the primary behaviors assessed with CBM, as with any assessment device, lack usefulness for *all* students. Work has proceeded with other CBM measures of reading than oral reading fluency (e.g., maze) and written expression (Tindal & Parker, 1989).

The lack of attention to reliability and validity of the other CBA models may stem from their primary use in making instructional planning or Exploring Solutions decisions. Evolving out of teacher informal testing using curricular materials, the foremost criterion for their quality was the degree to which they matched instructional content (i.e., content validity). Some researchers (e.g., Messick, 1989) have argued that content validity is not a form of validity but is a test construction issue. We believe strongly that CBA advocates *must* go beyond content validity to support their measures' quality. To the degree to which decisions other than Exploring Solutions are made, we must provide evidence that a test is accurate (reliable) and measures what it says it measures (valid). A necessary precursor to technical adequacy is explicit specification of measurement procedures.

*Application of specific CBA-model strategies across age ranges.* The procedures within most CBA models currently are associated with specific age- or grade-level populations. For example, the strategies represented by Neisworth and Bagnato (1986) are used with preschoolers, whereas Gickling's measurement procedures have an elementary-grade focus. It seems worthy to consider expanding the measurement strategies associated with the philosophical underpinnings of each model to other populations. The tenets of CBM—frequent, repeated measurement of key student outcome variables in an academic area for evaluating intervention effects—would be very useful for preschool populations.

For example, the Primary Prevention of Early Academic Problems (PPEAP) project currently is exploring downward extensions of CBM procedures to the kindergarten and first-grade levels (Good, Kaminski,



Schwarz, & Doyle, 1990). For preschool populations in particular, measures are needed that provide an estimate of the slope of pupil progress and a basis for ongoing, sequential decision making, with frequent opportunities to revise evaluations of risk (MacMann et al., 1989).

### Use of the Measures in Decision Making

We propose that evaluation of the utility of CBA be conducted within the framework of the problem-solving decisions (e.g., Problem Identification, Evaluating Solutions) described earlier in this chapter. These decisions form one dimension of Figure 2. The second dimension is that of the specific school-aged population that is to be investigated, preschool, elementary, and secondary pupils. A third dimension is that of a particular CBA model.

Interpreting Figure 2 then, one can identify research questions in Problem Identification with elementary-aged students using Gickling's CBA-ID model or Evaluating Solutions with secondary-aged students using CBA-CR strategies.

*Research on problem identification and certification.* With elementary-aged pupils, we believe that research on the use of CBM strategies as a reliable method of problem identification and certification (Shinn, Tindal, & Stein, 1988) has been exhausted. No more studies are really needed to confirm that students placed in special education generally are the lowest performers in a curriculum compared to their local peers. Few, if any, problem-identification studies have been conducted at the secondary or preschool levels with CBM. No published studies have been conducted using other models of CBA for making these kinds of decisions. If problem identification continues to be seen as an area of priority (which, for the most part, we do not), then research using other models and populations other than elementary-aged students should be conducted.

*Research on exploring solutions.* The major use of nearly all CBA procedures has been on identifying the content of instructional interventions, the "what to teach" (Marston, 1989). The underlying premise is that better assessment data about what students can do and need to do will result in better learning. In a sense, then, CBA data are independent variables that should be demonstrated empirically to improve student outcomes. In many ways, the intervention-planning information provided by CBA is a treatment that can be tested by using a treatment-evaluation model (Deno, 1986). As just one example, Gickling and Thompson (1985) propose that if students are placed in

instructional-level material they will make progress. If students are not placed in instructional-level material (i.e., frustration- or independent-level material), they will not make as much progress. Although this conception has great intuitive appeal, we argue the need for data on the effects of Gickling's placement criteria and suggest that other criteria may work better. The contribution of instructional placement criteria using CBM strategies (e.g., Deno & Mirkin, 1977) also lends itself to empirical investigation. The types of interventions derived from CBA data are virtually limitless. Given the magnitude of instructional problems in schools, we believe great efforts are needed to determine how data can be used to increase the likelihood of implementing effective programs and decrease the likelihood of implementing ineffective programs.

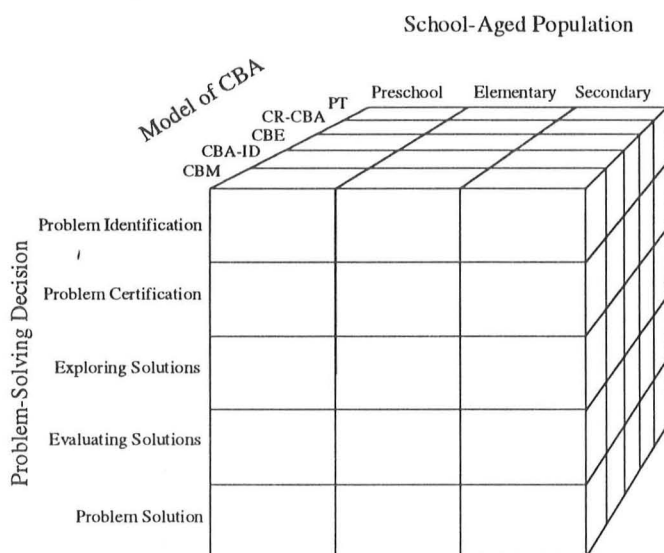


Figure 2. A matrix of research domains depicted by the type of problem-solving decision, type of CBA model, and target school-aged population.

A second key component of the Exploring Solutions decision is the specification of goals that are to be used to evaluate the effects of the intervention. Again, most of the research in using CBA to establish goals has been conducted within a CBM framework. The investigation of the effects of different goal structures and strategies on students' rates of progress and teaching (Fuchs & Fuchs, 1986; Fuchs, Fuchs, & Deno, 1985; Fuchs, Fuchs, & Hamlett, 1988; Fuchs, Hamlett, & Fuchs,

1988) has had fascinating outcomes. Some of the studies have been descriptive and need further experimental testing, however. For example, setting ambitious curricular goals has been associated with improved student outcomes (Fuchs et al., 1985). Other areas of research, such as the use of dynamic goals that change over time (Fuchs et al., 1988a), need replication.

*Research on evaluating solutions.* One of the most neglected decisions in schools is that of evaluating the effectiveness of interventions that are implemented. Far too often, *no* systematic data are collected to determine if what is implemented is working with individual students. When data are considered, they tend to be subjective opinions. Given the unique learning needs of individual students, as much or more assessment time and resources should be devoted to evaluating an intervention's effects as were used to identify the intervention's components. The evaluation of an intervention's effects using the curriculum in which students are instructed seems to be a logical process. Unfortunately, few *systematic* procedures for evaluating interventions using CBA have been specified in the professional literature. Even less research has been conducted in this area, with the exception of CBM. Within CBM, a host of research topics remain in making intervention effectiveness decisions. Among the important topics are further explication of the assets and liabilities of short-term versus long-term measurement with respect to estimating true progress, frequency of measurement, methods of summarizing student performance over time, and methods for increasing the frequency and effectiveness of changes in intervention strategies as a function of student performance data. Research on the use of computers in each of these areas (see Fuchs et al., 1988b, as well as this volume, for more details) also is increasing in prominence.

Efforts need to be increased on the use of other CBA strategies for evaluating interventions, in large part because CBM has been employed only to evaluate the effects of interventions in basic skill areas. Mastery monitoring approaches, where students' rates of progress through curricular objectives are examined (Jenkins, Deno, & Mirkin, 1979), remain potentially the most useful method in other curricular areas, especially for very young pupils and in secondary content areas. Unfortunately, mastery monitoring approaches have very few systematic procedures and virtually no research.

*Problem solution.* Problem solution decisions are made to determine if a problem is resolved and no longer requires additional resources. How do we know, for example, that an intervention has accomplished its purpose? In special education or Chapter I programs, this question

would be translated to mean, "What data do we have to suggest that special services are no longer required and a student may receive his/her instruction with other more typical students?" The use of data to make Problem Solution decisions is likely the least well-investigated area in education in general.

A problem-solving model would define a problem as resolved when the difference between what is expected and what occurs is no longer socially important. The use of student performance data in a curriculum again is logical for operationalizing what is expected and what is occurring and therefore may be useful in making this decision. No systematic procedures have been identified or developed, however. As a result, no empirical work has been accomplished, regardless of CBA model. Implementors of CBM (Allen, 1989; Shinn & Rodden-Nord, 1990) have begun a series of processes to assist educators in making Problem-Solution decisions.

### Research on Implementation

Most research on CBA strategies has been microcosmic, how specific measurement techniques work, and with what effects or how teachers can use specific decision rules to determine when to change their instructional programs. Very little research has been undertaken at a more molar, systems level, investigating, for example, what factors expedite or impede implementation. To date, the research that has been conducted has been constrained to CBM and from a retrospective perspective (Deno & Marston, 1989). Efforts should be made to study systems' reactions to implementation during the process of changes in assessment practices.

School district leaders (e.g., Germann, 1987) have identified a series of steps that are purported to increase the ease of implementation of CBM. If CBA is seen as a potential technology that should be implemented, then it seems logical that research on implementation should be conducted to facilitate the technology transfer. Prevailing opinion is that widespread changes occur neither easily nor frequently in education (Baer & Bushell, 1981; Cuban, 1990). Resistance factors should be identified and addressed.

CBA approaches, independently or in combination, represent innovations that will require change(s) in how schools operate. The assessment practices of school psychologists and special educators can be expected to change, as will the way the various service consumers (e.g., parents, teachers) accept and use the information that is provided. With reduced time spent on problem-identification and certification

decisions, it will be important to examine whether there are shifts in time devoted to intervention planning and evaluation of outcomes, and whether intervention services and resources can be restructured to serve students more effectively.

## CLOSING COMMENTS

CBA represents an important innovative assessment technology that has the potential to improve students' educational programs. We are pessimistic about whether the various CBA systems will be implemented with sufficient fidelity to improve outcomes, however. Although the appeal of using testing materials derived from students' curricula is obvious, we are of the opinion that the initial attraction may, in fact, be a distraction. That is, the use of content-valid tests is a necessary but not sufficient step for better educational assessment and decision-making practices. Just the use of content-valid tests stops at the superficial benefits of an alternative educational assessment approach. As we have illustrated, there is much more to improved educational assessment practices: A substantive shift in assessment paradigms is required. Through our examination of the literature written about CBA by its contributors and noncontributors, we believe that many knowledgeable persons are *not* seeing the required paradigmatic shift, and that what we will see is merely another test added to the repertoire of school psychologists and special educators. Better educational assessment practices cannot "combine state of the art regression discrepancy and curriculum-based models" (CASP, 1990, p. 12). Instructional plans derived from a profile analysis of WISC-R protocols are not well-wed to an analysis of CBA student error types.

Earlier, we pointed out Sarason's belief that school change comes when the system's values suggest that changes are necessary. We stated our own belief that leaders in school psychology have established a value system in which CBA may be integral. However, we are concerned that the "base of the triangle is not wide enough" to support the calls for changes in educational assessment practices espoused by CBA. That is, there may not be enough sufficiently trained personnel to implement quality educational assessment practices, including CBA, with sufficient integrity to change existing regularities. Training occurs at two levels, preservice and in-service. Bardon (1988) has pointed out the difficulties in training at both levels. The former requires training by institutions of higher education, which, as Bardon describes, are slow themselves to adopt new approaches. The difficulties of in-service training are compounded by the fact that many practitioners

consider themselves already trained and see little need for additional training, especially at the fundamental, conceptual level and to the degree that would be required by a major paradigm switch. For success, we will need to train well a *generation* of university trainers and school personnel. Changes in training programs may be occurring, but to date, changes in educational assessment training practices are not obvious (Reschly, Genshaft, & Binder, 1987).

Lest we close on a gloomy note, let us add that generally, schools that have implemented CBA-type procedures with integrity have reported positive outcomes (Germann & Tindal, 1985; Marston & Magnusson, 1985; 1988). Further, CBM is serving as an integral component of statewide adoption of a problem-solving assessment model and special education reform (Iowa State Department of Education, 1990).

In analyzing the characteristics of effectively implemented interventions described by Rogers (1983) (e.g., relative advantage, trialability, observability), we believe that each and all models of CBA possess many of these characteristics. The future of improved educational assessment using CBA strategies is filled with potential. We encourage a well-thought-out implementation process that exploits the limited technical assistance that is available.

## REFERENCES

- Alessi, G. (1988). Diagnosis diagnosed: A systemic reaction. *Professional School Psychology*, 3, 145-152.
- Alessi, G. (1988). Direct observation methods for emotional/behavior problems. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Behavioral assessment in schools: Conceptual foundations and practical applications* (pp. 14-75). New York: Guilford.
- Allen, D. (1989). Periodic and annual reviews and decisions to terminate special education services. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 184-203). New York: Guilford.
- Baer, D. M., & Bushell, D. (1981). The future of behavior analysis in the schools? Consider its recent past and then ask a different question. *School Psychology Review*, 10, 259-270.
- Bagnato, S.J., Neisworth, J.T., & Munson, S.M. (1989). *Linking developmental assessment and early intervention: Curriculum-based prescriptions* (2nd ed.). Rockville, MD: Aspen.
- Bailey, D. B. (1981). Investigation of learning measures as screening procedures with kindergartners. *Psychology in the Schools*, 18, 489-

495.

- Bardon, J. I. (1988). Alternative educational delivery approaches: Implications for school psychology. In J. L. Graden, J. E. Zins, & M. C. Curtis (Eds.), *Alternative educational delivery systems: Enhancing instructional options for all students* (pp. 563-571). Washington, DC: National Association of School Psychologists.
- Barlow, D. H., Hayes, S. C., & Nelson, R. O. (1984). *The scientist practitioner: Research and accountability in clinical and educational settings*. New York: Pergamon.
- Barrios, B. A. (1988). On the changing nature of behavioral assessment. In A. S. Bellack & M. Hersen (Eds.), *Behavioral assessment: A practical handbook* (3rd ed.) (pp. 3-41). New York: Pergamon.
- Blankenship, C. (1985). Using curriculum-based assessment data to make instructional decisions. *Exceptional Children*, 52, 233-238.
- Bursick, W. D., & Lessen, E. (1987). A classroom-based model for assessing students with learning disabilities. *Learning Disabilities Focus*, 3, 17-29.
- California Association of School Psychologists. (1990). *New directions: Registration packet for the California Association of School Psychologists 41st Annual Convention*. Millbrae, CA: Author.
- Carroll, J. B. (1989). The Carroll model: A 25-year retrospective and prospective view. *Educational Researcher*, 18 (1), 26-31.
- Coates, R. (1989). Faint praise. *NASP Communique*, 16, 3. Washington, DC: National Association of School Psychologists.
- Coulter, W. A. (1988). Curriculum-based assessment: What's in a name? *National Association of School Psychologists Communique*, 16, 13.
- Cuban, L. (1990). Reforming again, again, and again. *Educational Researcher*, 19 (1), 3-13.
- Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, 52, 219-232.
- Deno, S. L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review*, 15, 358-374.
- Deno, S. L. (1989). Curriculum-based measurement and alternative special education services: A fundamental and direct relationship. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 1-17). New York: Guilford.
- Deno, S. L., & Fuchs, L. S. (1987). Developing curriculum-based measurement systems for data-based special education problem solving. *Focus on Exceptional Children*, 19, 1-16.
- Deno, S. L., & Marston, D. (1989). Evaluation of CBM in public schools: Grant continuation proposal, U.S. Department of Education.



- Minneapolis, MN: University of Minnesota, Special Education.
- Deno, S. L., & Mirkin, P. K. (1977). *Data-based program modification: A manual*. Reston, VA: Council for Exceptional Children.
- Deno, S. L., Mirkin, P. K., & Shinn, M. R. (1979). *Behavioral perspectives on the assessment of learning disabled children* (Monograph No. 12). University of Minnesota, Institute for Research on Learning Disabilities.
- Deno, S. L., Mirkin, P. K., & Wesson, C. L. (1984). How to write effective data-based IEPs. *Teaching Exceptional Children*, 16, 99-104.
- Deno, S. L., Wesson, C., & King, R. P. (1984). Direct and frequent measurement of student performance: If it's so good for us, why don't we do it? *Learning Disability Quarterly*, 7, 45-48.
- Fuchs, L. S. (1986). Monitoring progress among mildly handicapped pupils: Review of current practice and research. *Remedial and Special Education*, 7, 5-12.
- Fuchs, L. S., Deno, S. L., & Marston, D. (1983). Improving the reliability of curriculum-based measures of academic skill for psychoeducational decisionmaking. *Diagnostic*, 8, 135-149.
- Fuchs, L. S., Deno, S. L., & Mirkin, P. K. (1983). Data-based program modification: A continuous evaluation system with computer software to facilitate implementation. *Journal of Special Education Technology*, 6, 50-57.
- Fuchs, L. S., Deno, S. L., & Mirkin, P. K. (1984). The effects of frequent curriculum based measurement and evaluation on pedagogy, student achievement and student awareness of learning. *American Educational Research Journal*, 21, 449-460.
- Fuchs, L. S., & Fuchs, D. (1984). Criterion-referenced assessment without measurement: How accurate for special education? *Remedial and Special Education*, 5, 29-32.
- Fuchs, L. S., & Fuchs, D. (1986a). Effects of systematic formative evaluation on student achievement: A meta-analysis. *Exceptional Children*, 53, 199-208.
- Fuchs, L. S., & Fuchs, D. (1986b). Curriculum-based assessment of progress towards long- and short-term goals. *The Journal of Special Education*, 20, 69-82.
- Fuchs, L. S., Fuchs, D., & Deno, S. L. (1982). Reliability and validity of curriculum-based informal reading inventories. *Reading Research Quarterly*, 18, 6-26.
- Fuchs, L. S., Fuchs, D., & Deno, S. L. (1985). The importance of goal ambitiousness and goal mastery to student achievement. *Exceptional Children*, 52, 63-71.
- Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1988). *Computer applications to*



- curriculum-based measurements: Effects of teacher feedback systems.* Nashville, TN: Peabody College, Vanderbilt University.
- Fuchs, L. S., Fuchs, D., & Tindal, G. (1986). Effects of mastery learning procedures on student achievement. *Journal of Educational Research*, 79, 286-291.
- Fuchs, L. S., Hamlett, C. L., & Fuchs, D. (1988). *Improving data-based instruction through computer technology: Description of Year 3 software.* Nashville, TN: Vanderbilt University.
- Galagan, J. E. (1985). Psychoeducational testing: Turn out the lights, the party's over. *Exceptional Children*, 52, 288-298.
- Germann, G. (1987, November). *Factors in CBM administration.* Presentation at the Minneapolis Special Education Leadership Conference, Minneapolis, MN.
- Germann, G., & Tindal, G. (1985). Applications of direct and repeated, curriculum-based assessment. *Exceptional Children*, 51, 110-121.
- Gickling, E., & Havertape, J. (1981). *Curriculum-based assessment (CBA).* Minneapolis, MN: National School Psychology In-Service Training Network.
- Gickling, E., & Thompson, V. (1985). A personal view of curriculum-based assessment. *Exceptional Children*, 52, 153-165.
- Good, R., Kaminski, R., Schwarz, I., & Doyle, C. (1990, April). *Using pupil progress to identify at-risk kindergarten students: Pilot study.* Paper presented at the 22nd Annual Convention of the National Association of School Psychologists, San Francisco, CA.
- Good, R. H., & Shinn, M. R. (1990). Forecasting accuracy of slope estimates for reading curriculum-based measurement: Empirical evidence. *Behavioral Assessment*, 12, 179-193.
- Graden, J. L., Zins, J. E., & Curtis, M. J. (1988). *Alternative educational delivery systems: Enhancing instructional options for all students.* Washington, DC: National Association of School Psychologists.
- Hall, G., & Hord, S. (1984). *Change in schools: Facilitating the process.* Albany, NY: State University of New York Press.
- Hargis, C. H. (1987). *Curriculum-based assessment: A primer.* Springfield, IL: Charles C. Thomas.
- Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1986). The treatment utility of assessment: A functional approach to evaluating assessment quality. *American Psychologist*, 42, 963-974.
- Heller, K. A., Holtzman, W., & Messick, S. (1982). *Placing children in special education: A strategy for equity.* Washington, DC: National Academy Press.

- Howell, K. (1986). Direct assessment of academic performance. *School Psychology Review*, 15, 324-335.
- Howell, K. W., & Kaplan, J. S. (1980). *Diagnosing basic skills: A handbook for deciding what to teach*. Columbus, OH: Charles Merrill.
- Howell, K.W., & Morehead, K. (1987). *Curriculum-based evaluation for special and remedial education*. Columbus, OH: Charles Merrill.
- Idol, L., Nevin, A., & Paolucci-Whitcomb, P. (1986). *Models of curriculum-based assessment*. Rockville, MD: Aspen.
- Iowa State Department of Education. (1990). *Problem-solving assessment series*: Des Moines, IA: Author.
- Jenkins, J. R., Deno, S. L., & Mirkin, P. K. (1979). Measuring pupil progress toward the least restrictive environment. *Learning Disability Quarterly*, 2, 81-92.
- Knutson, N., & Shinn, M. R. (in press). Curriculum-based measurement: Conceptual underpinnings and integration into problem-solving assessment. *The Journal of School Psychology*.
- Kratochwill, T. R., & Shapiro, E. S. (1988). Introduction: Conceptual foundations of behavioral assessment in schools. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Behavioral assessment in schools: Conceptual foundations and practical applications* (pp. 1-13). New York: Guilford.
- Lentz, F. E. (1988). Direct observation and measurement of academic skills: A conceptual review. In E. S. Shapiro & T. R. Kratochwill (Eds.), *Behavioral assessment in schools: Conceptual foundations and practical applications* (pp. 76-120). New York: Guilford.
- Lentz, F. E., & Shapiro, E. S. (1986). Functional assessment of the academic environment. *School Psychology Review*, 15, 346-357.
- Lombard, T. (1988a). Caution urged in embracing CBM assessment. *NASP Communique*, 20.
- Lombard, T. (1988b). *Curriculum-based measurement: Megatesting or mctestng*. St. Paul, MN: Minnesota State Department of Education.
- Macmann, G. M., Barnett, D. W., Lombard, T. J., Belton-Kocher, E., & Sharpe, M. N. (1989). On the actuarial classification of children: Fundamental studies of classification agreement. *The Journal of Special Education*, 23, 127-149.
- Marston, D. (1988). The effectiveness of special education: A time-series analysis of reading performance in regular and special education classrooms. *The Journal of Special Education*, 21, 13-26.
- Marston, D. (1989). Curriculum-based measurement: What is it and why do it? In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 18-78). New York: Guilford.
- Marston, D., Fuchs, L. S., & Deno, S. L. (1986). A comparison of standardized achievement tests and direct measurement techniques

- in measuring student progress. *Diagnostique*, 11, 77-90.
- Marston, D., & Magnusson, D. (1985). Implementing curriculum-based measurement in special and regular education settings. *Exceptional Children*, 52, 266-276.
- Marston, D., & Magnusson, D. (1988). Curriculum-based assessment: District-level implementation. In J. L. Graden, J. E. Zins, & M. C. Curtis (Eds.), *Alternative educational delivery systems: Enhancing instructional options for all students* (pp. 137-172). Washington, DC: National Association of School Psychologists.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: Macmillan.
- Mirkin, P. K., Deno, S. L., Tindal, G. A., & Kuehnle, K. (1982). Frequency of measurement and data utilization as factors in standardized behavior assessment of academic skill. *Behavioral Assessment*, 4, 361-370.
- National Association of School Psychologists/National Advocacy Coalition for Children. (1985). *Position statement: Advocacy for appropriate educational services for all children*. Washington, DC: Author.
- Neisworth, J. T., & Bagnato, S. J. (1986). Curriculum-based developmental assessment: Congruence of testing and teaching. *School Psychology Review*, 15, 266-287.
- Reschly, D. (1988). Special education reform: School psychology revolution. *School Psychology Review*, 17, 459-475.
- Reschly, D. J., Genshaft, J., & Binder, M. S. (1987). *The 1986 NASP survey: Comparison of practitioners, NASP leadership, and university faculty on key issues*. Washington, DC: National Association of School Psychologists.
- Reschly, D. J., Kicklighter, R., & McKee, P. (1988a). Recent placement litigation, Part I, Regular education grouping: Comparison of Marshall (1984, 1985) and Hobson (1967, 1969). *School Psychology Review*, 17, 9-21.
- Reschly, D. J., Kicklighter, R., & McKee, P. (1988b). Recent placement litigation, Part II, Minority EMR overrepresentation: Comparison of Larry P. (1979, 1984, 1986) with Marshall (1984, 1985) and S-1 (1986). *School Psychology Review*, 17, 22-49.
- Reschly, D. J., Kicklighter, R., & McKee, P. (1988c). Recent placement litigation, Part III, Analysis of differences in Larry P., Marshall, and S-1 and implications for future practices. *School Psychology Review*, 17, 39-50.
- Rogers, E. (1983). *Diffusion of innovations* (3rd ed.). New York: Free Press.
- Rosenfield, S., & Rubinson, F. (1985). Introducing curriculum-based

- assessment through consultation. *Exceptional Children*, 52, 282-287.
- Salvia, J., & Hughes, J. (1989). *Curriculum-based assessment*. New York: Prentice-Hall.
- Salvia, J., & Ysseldyke, J. E. (1987). *Assessment in special and remedial education* (4th ed.). Boston: Houghton Mifflin.
- Sarason, S. B. (1982). *The culture of the school and the problem of change* (2nd ed.). Boston: Allyn & Bacon.
- Shapiro, E. S., & Lentz, F. E. (1985). Assessing academic behavior: A behavioral approach. *School Psychology Review*, 14, 325-338.
- Shinn, M. R. (1986). Does anyone care what happens after the referral-test-placement process? The systematic evaluation of special education effectiveness. *School Psychology Review*, 15, 49-58.
- Shinn, M. R. (1988). Development of curriculum-based local norms for use in special education decision making. *School Psychology Review*, 17, 61-80.
- Shinn, M. R. (1989a). *Curriculum-based measurement: Assessing special children*. New York: Guilford.
- Shinn, M. R. (1989b). Identifying and defining academic problems: CBM screening and eligibility procedures. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 90-129). New York: Guilford.
- Shinn, M. R., Gleason, M. M., & Tindal, G. (1989). Varying the difficulty of testing materials: Implications for curriculum-based measurement. *The Journal of Special Education*, 23, 223-233.
- Shinn, M. R., Good, R. H., & Stein, S. (1989). Summarizing trends in student achievement: A comparison of evaluative models. *School Psychology Review*, 18(3), 356-370.
- Shinn, M. R., Goodwin, M., & Habedank, L. (1989). *Assessment of educational outcomes: Are we using the appropriate data to demonstrate effects?* Milwaukee, WI: Association of Applied Behavior Analysts.
- Shinn, M. R., & Marston, D. (1985). Differentiating mildly handicapped, low-achieving and regular education students: A curriculum-based approach. *Remedial and Special Education*, 6, 31-45.
- Shinn, M. R., Nolet, V., & Knutson, N. (1990). Best practices in curriculum-based measurement. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (2nd ed., pp. 287-308). Washington, DC: National Association of School Psychologists.
- Shinn, M. R., & Rodden-Nord, K. (1990). *A model of responsible reintegration of mildly handicapped students into general education classrooms*. Eugene, OR: University of Oregon.
- Shinn, M. R., Rosenfield, S., & Knutson, N. (1989). Curriculum-based assessment: A comparison of models. *School Psychology Review*, 18,

299-315.

- Shinn, M. R., Tindal, G., Spira, D., & Marston, D. (1987). Practice of learning disabilities as social policy. *Learning Disability Quarterly*, 10(1), 17-28.
- Shinn, M. R., Tindal, G., & Stein, S. (1988). Curriculum-based assessment and the identification of mildly handicapped students: A research review. *Professional School Psychology*, 3, 69-85.
- Shinn, M. R., Ysseldyke, J., Deno, S. L., & Tindal, G. (1986). A comparison of differences between students labeled learning disabled and low achieving on measures of classroom performance. *Journal of Learning Disabilities*, 19, 545-552.
- Singer, J. D., Palfrey, J. S., Butler, J. A., & Walker, D. K. (1989). Variation in special education classification across school districts: How does where you live affect what you are labeled? *American Educational Research Journal*, 26, 261-281.
- Taylor, R. L., Willits, P. P., & Richards, S. B. (1988). Curriculum-based assessment: Considerations and concerns. *Diagnostique*, 14, 14-21.
- Tindal, G. A., Fuchs, L. S., Fuchs, D., Shinn, M. R., Deno, S. L., & Germann, G. (1985). Empirical validation of criterion referenced tests. *Journal of Educational Research*, 78, 203-209.
- Tindal, G. A., & Parker, R. (1989). Development of written retell as a curriculum-based measurement in secondary programs. *School Psychology Review*, 18, 328-343.
- Tindal, G. A., Shinn, M. R., & Germann, G. (1987). The effect of different metrics on interpretations of change in program evaluation. *Remedial and Special Education*, 8, 14-28.
- Tucker, J. (1985). Curriculum-based assessment: An introduction. *Exceptional Children*, 52, 199-204.
- Wesson, C. L. (1987). Increasing efficiency. *Teaching Exceptional Children*, 20, 46-47.
- Wesson, C. L., King, R. P., & Deno, S. L. (1984). Direct and repeated measurement of student performance: If it's good for us, why don't we do it? *Learning Disability Quarterly*, 7, 45-48.
- Will, M. (1986). *Educating students with learning problems: A shared responsibility*. Washington, DC: U.S. Department of Education, Office of Special Education and Rehabilitative Services.
- Will, M. (1989). *The role of school psychology in providing services to all children*. Washington, DC: U.S. Department of Education, Office of Special Education and Rehabilitative Services.